T1Y1.1 /88-005
February, 1988

CONTRIBUTION TO: Telecommunications Standards Committee T1
Specialized Subjects Subcommittee T1Y1
Specialized Video and Audio Services Working Group T1Y1.1

SOURCE: National Telecommunications and Information Administration,
Institute for Telecommunication Sciences
Boulder, CO 80303
Edmund A. Quincy
Daniel J. Tomich
Robert F. Kubichek
(303) 497-5472

TITLE: **Expert Pattern Recognition Methodology for Technology Independent, User Oriented Classification of Video Transmission Quality**

ABSTRACT

This contribution proposes a new methodology for objectively determining video transmission quality that is user-oriented and independent of coding and transmission technologies. The technique employs an Expert Pattern Recognition System that operates on received image measurements with unknown transmission impairments. Quality classification by the system is provided on a minimum probability of error basis relative to subjective classification. The Expert System employs Statistical Pattern Recognition rules to perform low level processing and an initial classification. Additional knowledge about subjective classification, visual perception and user requirements are applied by the Expert System to obtain a refined classification of the image quality.

I. Introduction

A. Background

Existing standards for video transmission performance focus on traditional analog TV parameters such as waveform distortion and other measures as defined in EIA RS-170A, EIA RS-250B, and NTC Report No. 7. Standards development work has been motivated by an urgent need to characterize performance of existing transmission techniques, and has naturally focused on these well-understood traditional parameters.

This contribution proposes a methodology for developing user-oriented, technology-independent video transmission quality assessment using Expert Pattern Recognition (PR) techniques. Assessment would be accomplished by indirect comparison of the received video image with a standard transmitted source image. A small set of optimal parameters (features) would be measured on each image and their values used to represent images in the classification process. The system would be trained on the standard source image as well as representative distorted images during the development phase only. The training process develops the statistical parameters for the PR system as well as rules and facts for the Knowledge Base of the Expert System. During testing, prior knowledge of specific impairments introduced in transmission would not be required. Potential advantages are:

- Maximum provider flexibility in introducing technology enhancement. Technology dependent performance measures may necessarily restrict innovation, or provide an inadequate basis for evaluating proposed new systems.

- Ability to compare alternative systems (or system operating parameters) in user-oriented terms. This can improve provider efficiency and give users a practical means of comparing offered services.

- Ability to tailor service offerings to customer needs efficiently (e.g., premium quality grades or special military requirements).

These concepts have been applied to voice quality assessment by Quincy [1], and Quincy et al. [2,3]. Further, a Study Project Proposal being considered by T1Y1, suggests using this methodology for voice quality assessment [4].

B. Proposed Solution

This section discusses a possible test set design which will determine the end-to-end Quality of Service (QOS) of video transmission systems. The final product is envisioned as a standard source/receiver that replaces the video camera/monitor at each end of a transmission link. The test set would make the desired objective parameter measurements from each end and map the vector of parameter values into one of five classes of quality (overall goodness scale). An Expert Bayes Pattern Recognition System would perform the classification and can also account for other knowledge and data available including user bias. The end result would be a readout of one of these five classes (excellent, good, fair, poor or unsatisfactory), the associated confidence level and reasoning for the decision.

The salient features of this test system are:

- Maps objective measurements made on the output image into one of five classes
- Emulates subjective classification of distorted image data with minimum average probability of error (P_e) criterion
- Independent of system technology
- Interactive for user friendliness and adaption to user requirements
- Programmable

II. Proposed Expert Pattern Recognition Methodology

A. Expert System Structure

An expert system approach is selected for the classifier design in order to integrate all pertinent knowledge into the decision process, as suggested in Fig. 1. In order for a knowledge-based system to be an expert system [5, 6], it should supply the reasoning leading to its decision. The knowledge base will be constructed of facts and of rules based on expertise in pattern recognition, image processing, visual perception, and user requirements.

The Data Base (Fig. 1) serves as a dynamic scratchpad to store the sampled image obtained from the channel output, other specialized information about the channel and User, and parameter measurements. End-user requirements can be input to obtain tailored classifications for specific groups of users with such as those with color blindness or known transmission technologies. The Knowledge Base stores the domain expertise in the form of production rules (also referred to as situation - action or if - then rules). The Inference Engine uses a backward chaining strategy to apply production rules from the Knowledge Base to current measurements in the Data Base and obtain the best classification. On the average, the system solves the classification problem as would a subjective panel of viewers.

B. Objective Parameter Selection

The candidate objective video parameter set should include parameters from a variety of areas associated with image understanding and image processing. The candidate parameter list will include:

- Univariate parameters measured from the digitized image such as geometric distortion, spatial resolution, and spatial frequency measures.

- Parameters measured temporally, e.g. measured on luminance as a function of time. These could include Linear Predictive Coding (LPC) coefficients, Cepstral coefficients, etc.

- Bivariate or difference measures of brightness and color made between the undistorted image and the processed image.

Computer processing time and memory requirements for pattern recognition software grow as the number of parameters increases. Thus, it is advantageous to identify the smallest subset of candidate parameters that will provide sufficient discriminatory power for a Bayes minimum P_e classifier.

Most common methods of feature selection involve some suboptimum criterion associated with the feature space and do not deal directly with the ultimate goal of minimizing P_e. These methods usually focus on clustering effects of features with an attempt to maximize intraset distances while minimizing interset distances. However, they are not appropriate for complex multimodal data with non-linear decision boundaries.

We propose a feature selection algorithm that uses a bottom-up search to evaluate sets of candidate parameters [7]. The performance metric described in [7] provides nearly minimum P_e performance for a multimodal data set while keeping the problem tractable.

C. Statistical Pattern Recognition Classifier Module

Design of the classifier is founded on the requirement that it emulate subjective classification of distorted video with a minimum average probability of error criterion. This criterion is a special case of minimizing the average risk or loss as is done with the Bayes classifier.

In the general case for M classes the Bayes classifier computes a weighted likelihood from a measured parameter vector X as

$$d_i`(X) = p(\omega_i)L_i \; p(X|\omega_i) \quad , \quad i = 1, 2, \ldots, M. \tag{1}$$

for each of the M classes. Then it decides in favor of the class with the largest weighted likelihood $d_i`(X)$. In (1) $p(\omega_i)$ is the a priori probability of class ω_i, L_i is the loss associated with a decision in favor of ω_i and $p(X|\omega_i)$ is the conditional density of the parameter vector values for class ω_i. For minimum P_e, the losses L_i would all be equal for each class and can be deleted from (1) leaving just the joint likelihood of the vector parameter X being measured. The simplified classifier then selects the class corresponding to the largest of these likelihoods

$$d_i(X) = p(\omega_i) \; p(X|\omega_i), \quad i = 1, 2, \ldots, M. \tag{2}$$

Representative training data can be used to estimate (2) where $p(\omega_i)$ can be estimated as the ratio of outcomes in class i to the total number of outcomes. The joint parameter class density, $p(X|\omega_i)$, can be estimated as described in [3]. The probability or certainty of each quality class given the vector X of observed image parameters is obtained by (2) and Bayes rule as

$$p(\omega_i|X) = \frac{d_i(X)}{p(X)} = \frac{d_i(X)}{\sum_{j=1}^{5} p(X|\omega_j)P(\omega_j)}, \quad i = 1, 2, \ldots, 5. \tag{3}$$

A simple, efficient method for estimating each class probability density function (pdf), $p(X|\omega_i)$, from training data is discussed in [1].

III. Research and Development Steps

The test set design problem can be cast as an expert system problem with a Bayes pattern recognition classifier sub-module. Consequently, the procedure here reflects this type of system. The steps necessary to complete a design for the test set would include the following:

(1) Identify or develop a representative video image source for testing systems
(2) Develop a distorted video image data base from a variety of transmission techniques under various distorting conditions
(3) Develop training data by subjective classification of the distorted video image data base
(4) Objective parameter evaluation and selection
(5) Estimate statistics for each class from selected parameter set measurements on training data
(6) Design Bayes pattern recognition classifier
(7) Develop additional knowledge bases about subjective classification, video characteristics, visual impairments, visual perception, etc.
(8) Design the exert system classifier structure
(9) Validate the expert classifier

References

[1] E. A. Quincy, "Prolog-Based Expert Pattern Recognition Shell for Technology Independent, User-Oriented Classification of Voice Transmission Quality," Proceedings IEEE-ICC 87, Vol. 2, pp. 1164-1171, June, 1987.

[2] E. A. Quincy, N. B. Seitz, and R. F. Linfield, "Expert pattern recognition system for technology independent, user-oriented classification of voice transmission quality," ANSI Accredited Standards Working Group T1Q1.1, Contribution no. T1Q1.1/86-011, San Antonio, TX, Feb. 1986.

[3] E. A. Quincy, "Technology Independent Expert Pattern Recognition System for User-Oriented Classification of Voice Transmission Quality," ANSI accredited Standards Working Group T1Y1.2, Contribution no. T1Y1.2/87-044, Washington, DC, May 1987.

[4] E. A. Quincy and R. F. Kubichek, "Technology-Independent, User-Oriented, Objective Classification of Voice Transmission Quality," ANSI accredited Standards Working Group T1Y1.2, Contribution no. T1Y1.2/87-100, Research Triangle Park, NC, November, 1987.

[5] W. B. Gervator, Intelligent Machines, Prentice Hall, Englewood Cliffs, NJ, 1985.

[6] D. A. Waterman, A Guide to Expert Systems, Reading, MA: Addison Wesley, 1986.

[7] R. F. Kubichek and E. A. Quincy, "Statistical modeling and feature selection for seismic pattern recognition," Pattern Recognition, Special Issue on Geophysics, vol. 18, no. 6, pp. 441-448, 1985.

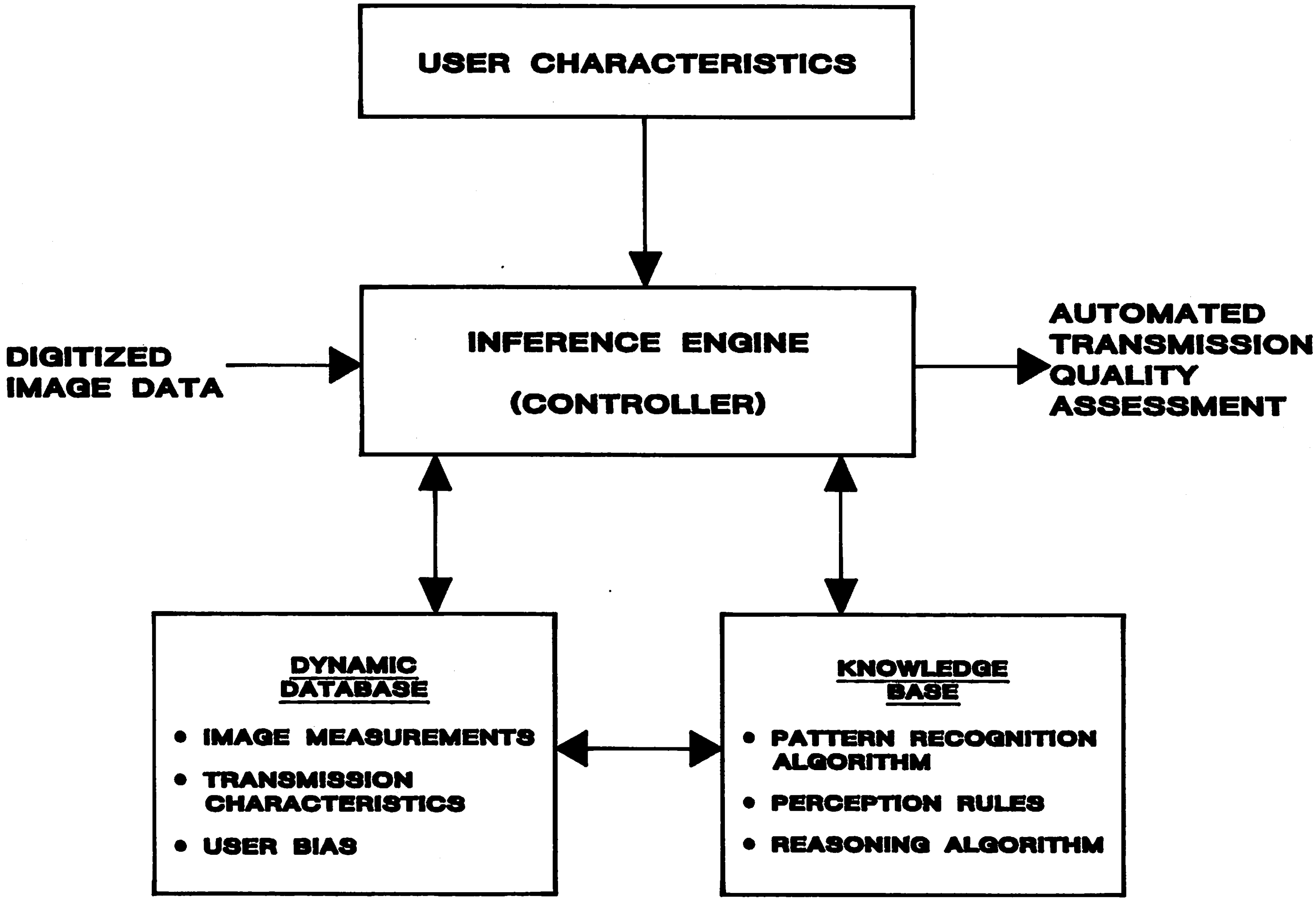

Fig. 1. Expert Systems Structure